

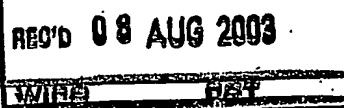
日本国特許庁  
JAPAN PATENT OFFICE

23.06.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日  
Date of Application: 2002年11月19日



出願番号  
Application Number: 特願 2002-334831  
[ST. 10/C]: [JP 2002-334831]

出願人  
Applicant(s): 松下電器産業株式会社

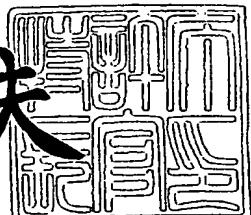
PRIORITY  
DOCUMENT

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2003年 7月25日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



BEST AVAILABLE COPY

出証番号 出証特 2003-3059215

【書類名】 特許願  
【整理番号】 2054041320  
【提出日】 平成14年11月19日  
【あて先】 特許庁長官殿  
【国際特許分類】 H04N 7/08  
【発明者】  
【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式  
会社内  
【氏名】 小林 正明  
【発明者】  
【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式  
会社内  
【氏名】 酒井 啓行  
【発明者】  
【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式  
会社内  
【氏名】 松井 謙二  
【発明者】  
【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式  
会社内  
【氏名】 桑野 裕康  
【特許出願人】  
【識別番号】 000005821  
【氏名又は名称】 松下電器産業株式会社  
【代理人】  
【識別番号】 100097445  
【弁理士】  
【氏名又は名称】 岩橋 文雄

## 【選任した代理人】

【識別番号】 100103355

## 【弁理士】

【氏名又は名称】 坂口 智康

## 【選任した代理人】

【識別番号】 100109667

## 【弁理士】

【氏名又は名称】 内藤 浩樹

## 【先の出願に基づく優先権主張】

【出願番号】 特願2002-182506

【出願日】 平成14年 6月24日

## 【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 メタデータ制作装置及び制作方法

【特許請求の範囲】

【請求項1】 コンテンツに関連するメタデータの制作装置であって、音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置。

【請求項2】 コンテンツに関連するメタデータの制作装置であって、音声入力手段と、音声認識手段と、キーボードを含むコンピュータ手段と、メタデータ制作手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換するとともに、上記認識されたデータが誤認識と判断された場合には、上記キーボードを含むコンピュータ手段により修正することを特徴とするメタデータ制作装置。

【請求項3】 コンテンツに関連するメタデータの制作装置であって、音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置。

【請求項4】 コンテンツに関連するメタデータの制作装置であって、音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて単語単位で認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置。

【請求項5】 コンテンツに関連するメタデータの制作装置であって、音声入力手段と音声認識手段とキーボードを含むコンピュータ手段とメタデータ

タ制作手段と上記コンテンツに関連した辞書とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて単語単位で認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換するとともに、上記認識されたデータが誤認識と判断された場合には、上記キーボードを含むコンピュータ手段により修正されることを特徴とするメタデータ制作装置。

**【請求項6】** 上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を備え、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1，2，3，4または5のいずれかに記載のメタデータ制作装置。

**【請求項7】** 上記コンテンツに付与されているコンテンツの番地あるいは番号あるいはフレームナンバーと上記メタデータとを入力し、コンテンツの番地あるいは番号あるいはフレームナンバー付メタデータを生成するコンテンツの番地あるいは番号あるいはフレームナンバー付与手段を備え、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1，2，3，4，5または6のいずれかに記載のメタデータ制作装置。

**【請求項8】** 上記コンテンツに関連した情報を上記音声入力手段により入力するにあたり、コンテンツをモニタする映像モニタ手段および音声モニタ手段のいずれかあるいは両方を確認しながら入力することを特徴とする請求項1，2，3，4，5，6あるいは7のいずれかに記載のメタデータ制作装置。

**【請求項9】** 上記コンテンツに関連した情報を上記音声入力手段により入力するにあたり、コンテンツをモニタする映像モニタ手段および音声モニタ手段のいずれかあるいは両方を確認し、かつ、コンテンツのシナリオまたはナレーション原稿を参照しながら入力することを特徴とする請求項1，2，3，4，5，6，7あるいは8のいずれかに記載のメタデータ制作装置。

**【請求項10】** コンテンツに関連するメタデータの制作方法であって、音声入力手段と、音声認識手段と、メタデータ制作手段とを用いて、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ制作手段によりメタデータに変

換することを特徴とするメタデータ制作方法。

【請求項11】 コンテンツに関連するメタデータの制作方法であって、音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを用いて、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて認識し、上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作方法。

【請求項12】 上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を用いて、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1または2のいずれかに記載のメタデータ制作方法。

#### 【発明の詳細な説明】

##### 【0001】

###### 【発明の属する技術分野】

本発明は、コンテンツ制作におけるメタデータ制作システム及び方法に関するものである。

##### 【0002】

###### 【従来技術】

近年、映像・音声コンテンツの制作において、これらコンテンツに関連したメタデータの付与することがおこなわれている。

##### 【0003】

しかしながら、上記メタデータの付与は、制作された映像・音声コンテンツのシナリオあるいはナレーション原稿をもとに、制作された映像・音声コンテンツを再生しながらメタデータとすべき情報を確認し、手作業でコンピュータ入力することにより制作する方法が一般的であり、相当な労力の必要な方法であった。

##### 【0004】

###### 【特許文献1】

特開平09-130736号公報

##### 【0005】

**【発明が解決しようとする課題】**

本願発明は、上記従来の問題点に係る課題を解決することを目的とするものであって、制作された映像・音声コンテンツを再生することによりメタデータとすべき情報を確認し、音声入力でコンピュータに入力することにより制作するシステム及び方法を提供することを目的とする。

**【0006】****【課題を解決するための手段】**

上記課題を解決するために本願発明は、制作されたコンテンツの再生手段、上記再生手段で再生された映像信号を表示する映像モニタ手段、上記再生手段で再生された音声信号をモニタする音声モニタ手段、上記映像モニタ手段および上記音声モニタ手段でオペレータが確認した制作すべきメタデータ内容をオペレータの発声によりマイクで収録する音声入力手段、上記音声入力手段により入力された音声信号を認識する音声認識手段、音声認識手段で認識された音声情報をメタデータに変換することによりメタデータを生成するメタデータ生成手段、および上記コンテンツと上記メタデータとを関連づけるため、上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力しタイムコード付メタデータとするタイムコード付与手段とを備えたものである。

**【0007】**

これにより、従来キーボードで入力し、制作していたメタデータを、音声認識を用いて音声入力し、自動的にタイムコード付きのメタデータを制作することができる。

**【0008】****【発明の実施の形態】**

本発明の請求項1に係る発明は、コンテンツに関するメタデータの制作装置であって、音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記コンテンツに関する情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置である。

。

**【0009】**

本発明の請求項2に係る発明は、コンテンツに関するメタデータの制作装置であって、音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関する辞書とを具備し、上記コンテンツに関する情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関する辞書に関連付けて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置である。

**【0010】**

本発明の請求項3に係る発明は、上記請求項1ないし上記請求項2に係る発明のメタデータ制作装置であって、上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を備え、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1または2のいずれかに記載のメタデータ制作装置である。

**【0011】**

以下、本発明の実施の形態について図面を用いて説明する。

**【0012】****(実施の形態1)**

図1は、本発明の実施の形態1によるメタデータ制作装置の構成を示すプロック図である。図1において、1はコンテンツ再生手段、2は映像モニタ手段、3は音声モニタ手段、4はマイク、5は音声認識手段、6はメタデータ生成手段、7はタイムコード付与手段、8は辞書である。

コンテンツ再生手段1は、たとえばVTR(あるいはハードディスクで構成された映像・音声信号再生装置、あるいは半導体メモリなどのメモリ手段を記録媒体とする映像・音声信号再生装置、あるいは光学記録式または磁気記録式などの回転型ディスクで構成された映像・音声信号再生装置、さらには伝送されてきたあるいは放送されてきた映像・音声信号を再生する映像・音声再生装置などのコンテンツ再生手段)である。上記コンテンツ再生手段1は、映像信号出力端子101、音声信号出力端子102およびタイムコード出力端子103を具備し、再生

された映像信号は端子101および201を介して、映像モニタ手段2に供給され、再生された音声信号は端子102および302を介して、音声モニタ3に供給され、再生されたタイムコードは端子103および703を介してタイムコード付与手段7に供給される。メタデータを制作する制作者（図示せず）は、映像モニタ手段2と音声モニタ手段3のいずれかあるいは両方を確認しながら、場合によればシナリオまたはナレーション原稿なども参照しながら、入力すべきメタデータを発声する。マイク4は、上記制作者の発声を受付、音声信号に変換して、音声認識手段5に供給する。また、必要に応じて、音声認識用の辞書8も、音声認識手段5に供給される。音声認識手段5で認識された、音声データはメタデータ生成手段6に供給され、メタデータあるいはタグに変換される。此のようにして、生成されたメタデータあるいはタグは、コンテンツ自身の内容と時間関係あるいはシーンとの関係を略略一致させるため、タイムコード付与手段7にて、コンテンツ再生手段1から供給されたタイムコード情報が付与される。

### 【0013】

より具体的に説明するため料理説明をする場面を一例として想定する。この場合オペレータが、映像モニタ手段2の表示画面を確認しながらマイク4を通じて”塩 1さじ”と発生すると、音声認識手段5にて、辞書8を参照しながら、”塩”、”1さじ”と認識されメタデータ生成手段6にて各々”塩”、”1さじ”というタグに変換される。なお、音声認識としては、上記音声認識手段5に限定されず、一般的に用いられている種々の手段を用いて音声認識し”しお”、”ひとさじ”とのデータを認識してもよい。

なお、一般には、メタデータとは、このようなタグの集合体を意味するものである。タイムコード付与手段703では、コンテンツ再生手段103からの信号をもとに、タイムコードが付与されたタイムコード付与メタデータが生成される。具体的には、図2に示すようなパケットデータが生成される。生成されたメタデータは、そのまま出力されても良く、またハードディスク等の記録媒体に蓄積しても良い。また、上記の実施例においては、パケット形式のメタデータを生成するとして説明したが、特に限定されるものではない。

### 【0014】

さらに上述した実施例では、コンテンツとしてタイムコードの付与されている動画コンテンツの場合について説明したが、静止画コンテンツあるいは、デジタルデータコンテンツなどの場合には、上記静止画コンテンツあるいはデジタルデータコンテンツを識別するために、動画の場合のタイムコードに相当するコンテンツの番地あるいは番号を用いて上記コンテンツと生成された上記メタデータを関連づけてもよい。

#### 【0015】

なお、一般的には、音声認識には何らかの誤認識が生じる可能性がある。誤認識が生じた場合、制作されたメタデータ、タグをコンピュータ手段などの情報処理手段を用いて修正することは可能である。

#### 【0016】

##### 【発明の効果】

以上説明したように発明は、コンテンツに関連したメタデータの作成あるいはタグ付けを行うに当たり、音声入力による音声認識を用い、且つ、上記メタデータあるいはタグとコンテンツとの時間あるいはシーンとの関連付けを行うため、従来のようなキーボード入力より、効率的に、メタデータの作成やタグ付けを実施することができる。

##### 【図面の簡単な説明】

###### 【図1】

本発明の実施の形態1に係るメタデータ制作装置の構成を示すブロック図

###### 【図2】

本発明の実施形態1に係るタイムコード付きメタデータの一例を示す図

##### 【符号の説明】

- 1 コンテンツ再生手段
- 2 映像モニタ手段
- 3 音声モニタ手段
- 4 マイク
- 5 音声認識手段
- 6 メタデータ生成手段

7 タイムコード付与手段

8 辞書

101 映像出力端子

102 音声出力端子

103 タイムコード出力端子

201 映像入力端子

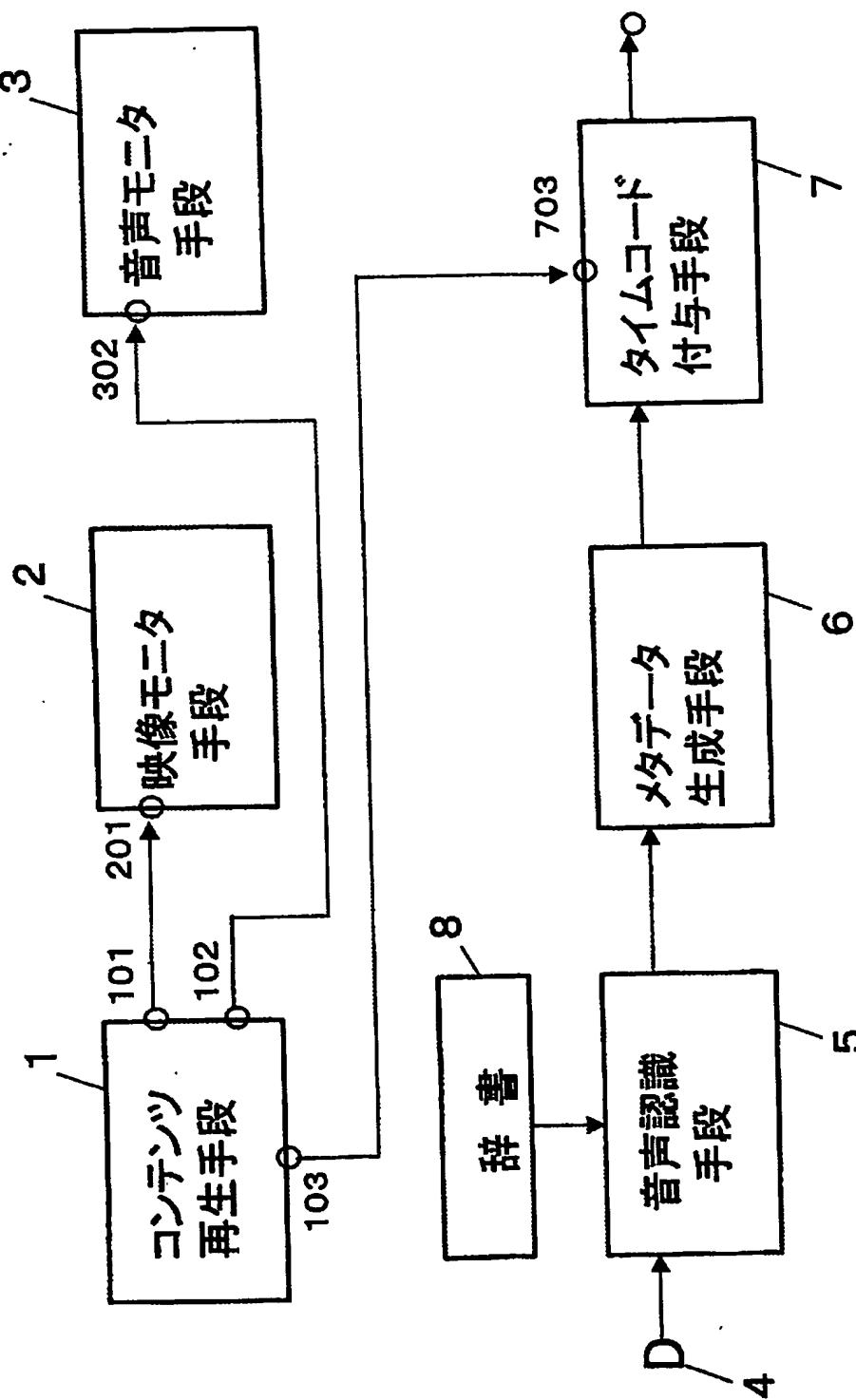
302 音声入力端子

703 タイムコード入力端子

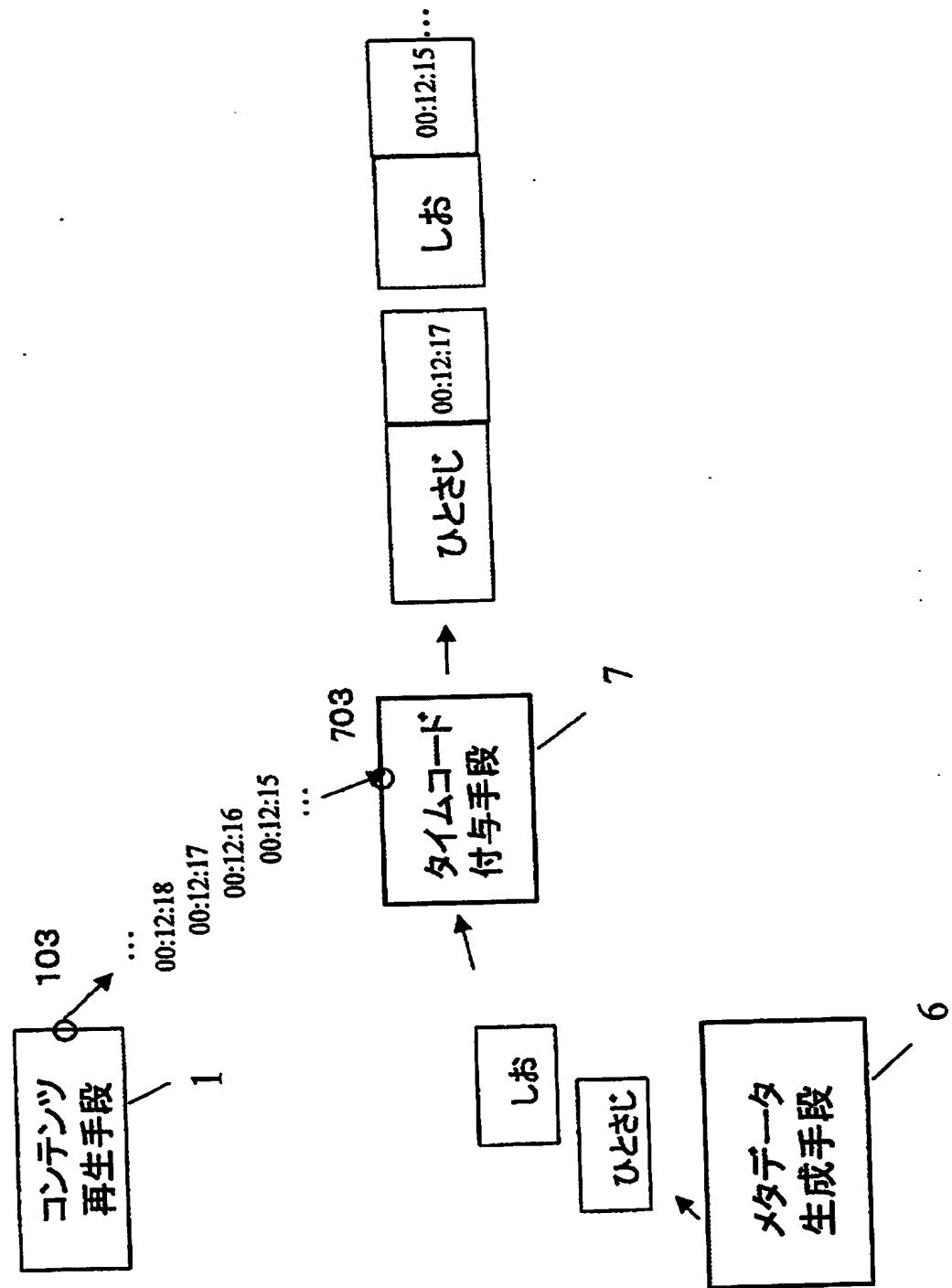
【書類名】

図面

【図1】



【図2】



【書類名】 要約書

【要約】

【課題】 映像・音声コンテンツの制作において、これらコンテンツに関連したメタデータの付与するに際し、付与する際の労力を低減することを目的とする。

【解決手段】 音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ制作手段によりメタデータに変換することにより、自動的にメタデータ、タグとコンテンツとの時間あるいはシーンとの関連付けを行う。

【選択図】 図1

特願2002-334831

出願人履歴情報

識別番号 [000005821]

1. 変更年月日 1990年 8月28日

[変更理由] 新規登録

住所 大阪府門真市大字門真1006番地  
氏名 松下電器産業株式会社